Unclassified

(2)

AD-A216 496

RT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION: Unclassified	1b RESTRICTIVE MARKINGS: N/A
2a SECURITY CLASSIFICATION AUTHORITY: N/A	3 DISTRIBUTION/AVAILABILITY OF REPORT: Distribution unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): 3-5-25038	5 MONITORING ORGANIZATION REPORT NUMBER(S): N/A
6a NAME OF PERFORMING ORGANIZATION: Baylor College of Medicine; 6b OFFICE SYMBOL (If applicable): N/A	7a NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c ADDRESS (City, State, and ZIP Code): Center for Biotechnology, 4000 Research Forest Drive, The Woodlands, TX 77381	7b ADDRESS (City, State, and ZIP Code): 800 N. Quincy St., Arlington, VA 22217-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b OFFICE SYMBOL (If applicable): ONR	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-89-J-1167
8c ADDRESS (City, State, and ZIP Code): 800 N. Quincy St., Arlington, VA 22217-5000	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61153N; PROJECT NO.; TASK NO 441-N012-01; WORK UNIT ACCESSION NO.

11 TITLE (Include Security Classification)

The Design of Oligonucleotides which Attack Specific Gene Targets

12 PERSONAL AUTHOR(S)
Hogan, Michael E.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Annual	FROM 12/88 TO 11/89	1989, December 8	19

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	DNA Recognition, Triplet helix formation,
06	03		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

During the first year of Navy support, we have refined our understanding of triple helix formation and in the process, have synthesized a set of triplex forming oligonucleotides directed against sites in the control region of interesting eukaryotic genes. We have shown that, in favorable instances, the dissociation constant for triplex formation can exceed 10-9M under physiological conditions. Work has begun to extend those structural analyses, by application of molecular modeling and 2D NMR methods. Work has also begun to apply photochemical analysis of triplex formation for the purpose of mapping triplex formation on gene target sites (HPRT and EGFR) within intact eukaryotic cells.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	(U)
22a NAME OF RESPONSIBLE INDIVIDUAL: Dr. Michael Marron	22b TELEPHONE (Include Area Code): 202/696-4760; 22c OFFICE SYMBOL: ONR

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

ANNUAL REPORT DATE: 21 NOVEMBER, 1989

CONTRACT N00014-89-J-1167, R&T CODE 441 N012-01

PRINCIPAL INVESTIGATOR: Dr. Michael E. Hogan

CONTRACTOR: Baylor College of Medicine

CONTRACT TITLE: The Design of Oligonucleotides Which Attack Specific Gene Targets

START DATE: DECEMBER 1988

RESEARCH OBJECTIVE: To study the site selectivity of triplex-forming oligonucleotides (TFOs) under physiological conditions of salt and pH, employing a combination of molecular modeling and physical experimentation to define the rules which determine site selective oligonucleotide binding. That understanding will be used to design molecules with the capacity to bind to discrete gene targets in the nucleus of an intact cell.

SUMMARY OF WORK PLAN FOR THE NEXT YEAR OF SUPPORT:

Based upon molecular modeling, chemical and biochemical analysis and NMR, our plan is to refine the rules of triplex formation (including those with modified bases). We are particularly interested to evaluate the role of nearest neighbors: ie, to consider target sites in terms of three base pair segments and to begin to evaluate the efficacy of non-standard nucleotide substituents.

The second goal is to begin to use TFOs as reagents to induce site-selective gene destruction at two DNA targets: the HPRT gene and the control region of the human epidermal growth factor receptor (EGFR).

The HPRT gene was chosen because it is X linked, and therefore haploid in the (male) cells under study. After cells are treated with the eosin conjugate and irradiated (our AR+ laser), mutant cells will analyzed by the standard positive HPRT selection methods. The DNA region of interest in the resulting HPRT- mutants will be expanded and sequenced by PCR, to confirm that damage has been targeted to the site of triplex formation.

The EGFR gene has been chosen because its promotor domain assumes a DNaseI hypersensitive configuration which is associated with genes which have been activated for the purpose of transcription. It has also been chosen because the functional state of the gene is easy to assess in a cultured cell line (A431 carcinoma).

SUMMARY OF RECENT RESEARCH PROGRESS:

STRUCTURE AND STABILITY DETERMINATION IN VITRO.

We have synthesized a set of triplex-forming oligonucleotides and

a complementary set of duplex target sites for the purpose of defining the base composition dependance of triplex formation. Some of those data are presented at the end of this section. Triplex formation has been studied by the band shift method, DNaseI footprinting and most recently by direct photochemical mapping of dye-oligonucleotide conjugates. For the purpose of the photochemical analysis, we have modified published procedures for affixing primary amine linkers to the 3' and 5' termini of oligonucleotide reagents, which we use as the site of attachment of isothiocyanate moieties: the isothiocyanates of eosin, fluorescein, pyrene and tetramethyl rhodamine. A new dye conjugate chemistry, which promises to be superior to these, is under investigation (see below).

As described in the biannual report, based upon those initial studies, we have significantly refined our understanding of triplex formation. Below, we list our current understanding of the "rules".

Rule 1. Define a binding site domain. In general, the stability of triplex formation increases monotonically with the length of the binding domain. At present, binding sites in the 25-35 base pair range appear to be optimal, base upon consideration of cell uptake, stability and synthetic considerations.

Rule 2. Define the gross purine/pyrimidine symmetry of the binding site. We define the more purine rich of the two strands of the binding site as the "orienting" strand. As seen from preliminary data in our laboratory (at the end), triplex formation is most stable on targets where the orienting strand is purine rich: but the rate of stability decrease with respect to decreasing purine content is relatively modest. Several of the targets we have studied fall into a target class which is best described as having an orienting strand with a "purine rich" end abutting a "pyrimidine rich" end (ie, the IL2R target). In the context of our current understanding, the triplex which results from oligonucleotide binding to such sites may develop a structural discontinuity at the border between the two sub-domains within the target sequence.

Rule 3. Previous work had suggested that a TFO is best designed so that it can bind throughout the major groove of the target with its backbone parallel to that of the orienting strand. However, during the past year we have shown that the preference may not be large and may itself be dependent on sequence details of the target. Several examples of that novel effect are displayed below as preliminary results.

Rule 4. In either the parallel or antiparallel orientation, place G in the triplex-forming oligonucleotide in opposition to a GC base pair in the target (thereby forming a GGC triplet) and a T in opposition to an AT base pair (to form the TAT) triplet. Such local triplet bonding schemes give rise to H bonding between the triplex forming oligonucleotide and the purine of the underlying Watson Crick base pair. Such bonding is likely to be of the Hoogsteen or reverse Hoogsteen type, depending on sequence context; ie G•G and T•A bonding respectively. The plan during year two is to use 2D NMR

methods to determine if that degree of conformational freedom is actually being employed.

Formation of such "external" bonds occurs through the major groove of the helix, and does not significantly alter the underlying target duplex. Work is in progress to determine if there are target sequences which produce exceptions to the above simple G=GC and T=AT rule (such as the junctional domain of a bifurcated target site, see above). Work is also in progress to determine if chemical moieties can be found which are superior to T and G in terms of the capacity to H bond in this way.

For the c-myc target site (Fig. 1), the antiparallel triplex is preferred under all conditions tested thusfar: to the extent that, if a TFO is designed so as to bind parallel, the antiparallel complex will form at the expense of several mismatches in the bound complex.

For the interleukin 2 receptor alpha chain target, the distinction is less pronounced: both the parallel and antiparallel triplex have similar stability, but with the apparent dissociation constant for formation of the antiparallel triplex which is approximately ten times that of the TFO isomer designed to bind parallel.

Because the physical composition of the myc and IL2R targets are quite different, we are reasonably sure that the observed parallel/antiparallel distinction is not a simple property of GC content. Work is in progress to understand the origin of the orientational preference. Recent data of that kind, derived from triplex formation within the epidermal growth factor receptor and insulin receptor gene promotors (Fig.1) will be useful in that regard.

At present, we believe that this aspect of triplex formation is ambiguous enough that we have begun to synthesize both the parallel and antiparallel isomers to any interesting target.

In order to rationalize the observed equivalence of the parallel and antiparallel triplex, we have begun a comprehensive program of NMR and molecular modeling, in collaboration with Dr. B. Monte Pettitt of the University of Houston Institute for Molecular Design. The goal of this project is introduce selected distance information derived from NOESY spectroscopy into CHARM, thereby constraining the range of solutions to the structure refinement process. This program of NMR and modeling refinement is a major component of the work to be performed in the next year.

TFO BINDING WITHIN THE NUCLEUS

Our bias is that DNA in chromatin is a fundamentally different cellular target than is "naked DNA". Therefor, to understand to gene selectivity of TFO binding in a context which has functional importance, it is necessary to devise methods to study TFO binding selectivity in an intact nucleus. To achieve that end, we have chosen to synthesize TFO conjugates with DNA photocleavage reagents, so that, when paired with DNA hybridization methods, TFO binding in the nucleus can be assessed from the resulting pattern of laser induced DNA cleavage.

During the past year, we have established two test systems to begin analysis of gene-specific DNA damage in a living cell: one

to assess TFO-induced damage genetically within a coding region; the second to detect TFO-induced damage within a gene control region. We have determined that the human HPRT gene is well suited as a target for the genetic approach and in collaboration with Tom Wilson's laboratory (Baylor Biochemistry) we have begun to study triplex-induced photochemical damage within the first exon of the gene. To date, we have fabricated eosin-oligonucleotide conjugates (and fluorescein and rhodamine conjugates as controls), have measured their binding affinity to the exon 1 target site and have confirmed (by fluorescence microscopy) that these water soluble dyes form conjugates which are transported efficiently into the sarcoma cell line which is being used as a test system. Intracellular analysis of TFO binding to the HPRT exon 1 (as determined genetically) will begin after the first of the year.

To assess TFO binding within a promotor region, we have chosen to study the human EGFR gene because it assumes a "classical" DNaseI hypersensitive chromatin structure within a 200bp domain within has been determined to contain the prominent control elements for the gene. To date, we have obtained cloned DNA (pER9GEM3) corresponding to the promotor region from the laboratory of origin (Dr. Glenn Merlino NCI), have designed and synthesized mg quantities of TFOs to several sites in the EGFR promotor and have measured their site selectivity in vitro (Fig. 1). As seen in Fig. 1, the EGFR-specific TFOs display high affinity (dissociation constants in the nanomolar range) for their respective target sequence under more-less physiological conditions.

A341 carcinoma cells have been established in culture in the laboratory of our collaborator M.Duvic (U. of T Health Sciences Center) and are ready for TFO uptake analysis and for photochemical mapping, which will begin immediately. The goal of those experiments is to perfuse dye-TFO conjugates into A431 cells, irradiate in situ at 500nm, then extract DNA then analyze the resulting strand break distribution at 10 base resolution by hybridization methods. High resolution determination of TFO binding to the EGFR control region (as compared to its specificity with naked DNA) is a major goal of the work for the next year of support.

NEW CLASS OF DYE-TFO CONJUGATE

Very recently, we have begun to exploit what we believe will prove to be a much superior dye probe for such oligonucleotide coupling: a modified methyl propidium derivative (carboxy methyl propidium CMP), first described as a bisintercaltor by Dervan (1) then as an affinity chromatography reagent by the BPS group (2). As we have shown recently for ethidium (3), a close homologue of methyl propidium, the dye engages in an efficient, stoichiometric DNA cleavage reaction when irradiated at 500nm with an Ar+ ion laser.

At the end (Fig. 2), we describe the chemistry employed to affix CMP to triplex forming oligonucleotides and the DNA photocleavage properties of the dye. Like ethidium, the CMP moiety is useful in that in the uncoupled state, it is not transported into cells. Moreover, the dye displays little fluorescence until bound to a duplex DNA target. Together, those properties suggest

that fluorescence microscopy can be used to monitor the kinetics of CMP-oligonucleotide uptake into the nucleus.

A major scientific goal during the next year of support will be to use CMP-TFO conjugates as DNA photocleavage reagents and as a tools to monitor the mechanism of TFO uptake into the nucleus.

NON-STANDARD NUCLEOSIDE SUBSTITUENTS

Our bias is that the naturally occurring DNA bases may be a relatively poor solution to the triplex design problem. As described in the biannual report, we have contracted Glenn research Inc. to prepare 1 gram of the beta cyanoethyl phosphonate derivative of xanthine, based upon the suspicion that it may be much better than T as a component of a triplex-forming oligonucleotide. To date, they have not successfully completed the synthesis in sufficient yield. During the next year of support, the xanthine derivatives will be tested in the c-myc class of TFO, as the triplex complement to AT base pair sites in the duplex target.

INVENTIONS (YEAR 1): Baylor filed a patent in December, 1988 on the application of the triplex technology in a pharmaceutical context. Navy-funded work had not yet been performed.

PUBLICATIONS & REPORTS: Two papers are in preparation. A book chapter has been published and is included. Hogan recently presented some of the triplex work at the NCI conference: "Oligonucleotides as Antisense Inhibitors of Gene Expression: Therapeutic Implications", Rockville MD, 6/89. Hogan has also been asked to discuss triplex-forming oligonucleotides at the MD Anderson Cancer Center (January), Duke and University of North Carolina, Chapel Hill (March) and at the upcoming Jerusalem Symposium " Molecular Basis of Specificity in Nucleic Acid-DNA Interaction", Jerusalem, May, 1990.

TRAINING ACTIVITIES: One post doctoral fellow, Ross Durland, is supported from the contract.

LITERATURE CITED:

1. P.B. Dervan & M.M. Becker (1978) J.A.C.S. 100:1969-1970
2. J.D. Hardy etal (1989) Nucl. Acids Res. 17:6950-6955
3. G. Krishnamurthy et al (1989) Biochemistry, in the press.

Figure 1. Equilibrium Binding Data for Triplex Formation.

Data have been accumulated at 37C in a standard binding buffer, comprising: 20mM Tris/HCl, pH 7.8, 5mM MgCl, 5mM spermine. We have found that this assay system optimizes binding for TFOs designed as described above. It should be noted that this buffer is quite "physiological" in its composition in that most DNA restriction enzymes and polymerases retain high activity in this assay buffer. Below, data have been catalogued by gene target. In each instance, the dissociation constant has been measured by electrophoresis, monitoring the binding of radiolabelled TFO to unlabelled cloned DNA target sites (binding is manifest as co-migration of the TFO as the triplex through a gel matrix). Under those conditions, nonspecific binding to sequences other than the target has not been detected in any test system thus far.

I. THE C-MYC PROMOTOR

Inappropriately high levels of c-myc gene expression are strongly associated with the incidence of a variety of human tumors. Numbering of the duplex target site is relative to the principal mRNA start site.

```
 -153                                                          -99
5'-TCTCCTCCCCACCTTCCCCACCCTCCCCACCCTCCCCATAAGCGCCCCTCCCGGG-3'
3'-AGAGGAGGGGTGGAAGGGGTGGGAGGGGTGGGAGGGGTATTCGCGGGGAGGGCCC-5'
5'-GTGGTGGGGTGGTTGGGGTGGGTGGGGTGGGTGGGGT-3' GT37anti
```

Target function. The principal activating protein binding site of the c-myc gene promotor.

Kdiss = 5x10-10 molar in TFO equivalents.

II. THE EGFR PROMOTOR

Inappropriately high expression of the epidermal growth factor gene (EGFR) has been implicated as crucial to the development of cancers and several skin diseases (psoriasis). Numbering of the duplex target site is relative to the principal mRNA start site.

```
 -109                        -83
5'-TCCGCCGAGTCCCCGCCTCGCCGCC-3'
3'-AGGCGGCTCAGGGGCGGAGCGGCGG-5'
3'-TGGGGGGTGTGGGGGGGTGGGGGGG-5' EGFR1p
5'-TGGGGGGTGTGGGGGGGTGGGGGGG-3' EGFR1ap
```

Target function: SP1 binding site.

Kdiss > 1x10-6 molar: ie not detectible as yet, for either isomer up to 10-6M of added target duplex.

-307 -281
5'-TCCCTCCTCCTCCCGCCCTGCCTCCCC-3'
3'-AGGGAGGAGGAGGGCGGGACGGAGGGG-5'
3'-TGGGTGGTGGTGGGGGGGTGGGTGGGG-5' EGFR2p
5'-TGGGTGGTGGTGGGGGGGTGGGTGGGG-3' EGFR2ap
Target function: SP1 binding site.

EGFR2ap
Kdiss = 5x10-8 molar in TFO equivalents.

EGFR2p
Kdiss > 1x10-6 molar; ie, not detectible at 10-6M of target duplex.

-352 -318
5'-TTCTCCTCCTCCTCTGCTCCTCCCGATCCCTCCTCC-3'
3'-AAGAGGAGGAGGAGACGAGGAGGGCTAGGGAGGAGG-5'
3'-TTGTGGTGGTGGTGTGGTGGTGGGGTTGGGTGGTGG-5' EGFR3p
5'-TTGTGGTGGTGGTGTGGTGGTGGGGTTGGGTGGTGG-3' EGFR3ap
Target function: SP1 binding site

EGFR3ap
Kdiss = 1x10-8 molar in TFO equivalents.

EGFR3p
Kdiss > 1x10-6 molar; ie, not detectible at 10-6M of target duplex.

-363 -338
5'-TTCTCCTCCCTCCTCCTCGCATTCTCCTCCTCCTCT-3'
3'-AAGAGGAGGGAGGAGGAGCGTAAGAGGAGGAGGAGA-5'
3'-TTGTGGTGGGTGGTGGTGGGTGGGTGGTGGTGGTGT-5' EGFR4p
5'-TTGTGGTGGGTGGTGGTGGGTGGGTGGTGGTGGTGT-3' EGFR4ap
Target function: Nuclease sensitive domain required for EGFR expression.

EGFR3ap
Kdiss = 1x10-8 molar in TFO equivalents.

EGFR3p
Kdiss > 1x10-6 molar; ie, not detectible at 10-6M of target duplex.

III.MOUSE INSULIN RECEPTOR

The mouse insulin receptor has been chosen for study because of the similarity between its promotor domain and that of EGFR and because the function of the IR gene is particularly easy to study in vitro.

-200 -161
5'-GGGGGAAAGGGGCCGAGGAGGCGAGGAGGCGAGGAGGC-3' TARGET
3'-CCCCCTTTCCCCGGCTCCTCCGCTCCTCCGCTCCTCCG-5' SITE
3'-GGGGGTTTGGGGGGGTGGTGGGGTGGTGGGGTGGTGG-5' IR1ap
5'-GGGGGTTTGGGGGGGTGGTGGGGTGGTGGGGTGGTGG-3' IR1p
TARGET FUNCTION: Upstream region with similarity to SP1 domain in human allele.

IR1p
Kdiss = 1x10-8 molar in TFO equivalents.

IR1ap
Kdiss = 5x10-8 molar

```
 -137                                                  -99
5'-CCTCCCCGAGCTCTGCTGGCTTTCTCCCCCTCCCCCCCC-3' TARGET
3'-GGAGGGGCTCGAGACGACCGAAAGAGGGGGAGGGGGGGG-5' SITE
3'-GGTGGGGGTGGTGTGGTGGGTTTGTGGGGGTGGGGGGGG-5' IR2 par
5'-GGTGGGGGTGGTGTGGTGGGTTTGTGGGGGTGGGGGGGG-3' IR2 anti
```

TARGET FUNCTION: Upstream region with TC repeats, in human & mouse

IR1p

Kdiss = 1x10-7 molar in TFO equivalents.

IR1ap

Kdiss = 5x10-7 molar

IV. INTERLUKIN 2 RECEPTOR, ALPHA CHAIN

The interleukin 2 receptor is a crucial element in the T cell response to antigens. As such, reagents which interfere with the expression of the IL2R alpha subunit gene may have usefulness as antirejection agents: "son of cyclosporin". Below we list a family of TFO reagents which have been targeted against important sites within the control region of IL2R. Numbering of the duplex target site is relative to the principal mRNA start site.

```
 -273                                                  -224
5'-CAACGGCAGGGGAATCTCCCTCTCCTTTTATGGGCGTAGTGAAGAAAGGA-3'
3'-GTTGCCGTCCCCTTAGAGGGAGAGGAAAATACCCGCATCACTTCTTTCCT-5'
               3'-TGTGGGTGTGGTTTTTTGGGGGTTGTGTTGTTTGGT-5' IL36par
               5'-TGTGGGTGTGGTTTTTTGGGGGTTGTGTTGTTTGGT-3' IL36anti
 3'-TTGGGGTGGGGTTTGTGGGTGTGGTTTT-5' IL28par
 5'-TTGGGGTGGGGTTTGTGGGTGTGGTTTT-3' IL28anti
 3'-TTGGGGTGGGGTTTGTGGGTGTGGTTTTTTGGGGG-5' IL35par
 5'-TTGGGGTGGGGTTTGTGGGTGTGGTTTTTTGGGGG-3' IL35anti
```

TARGET FUNCTION:The NF Kappa Beta protein factor binding site, which is the principal tissue specific activator of the alpha chain of IL2R. Several solutions have been listed, to account for the parallel/antiparallel homology and to account for uncertainty as to the breadth of the NFKB binding domain.

IL28ap

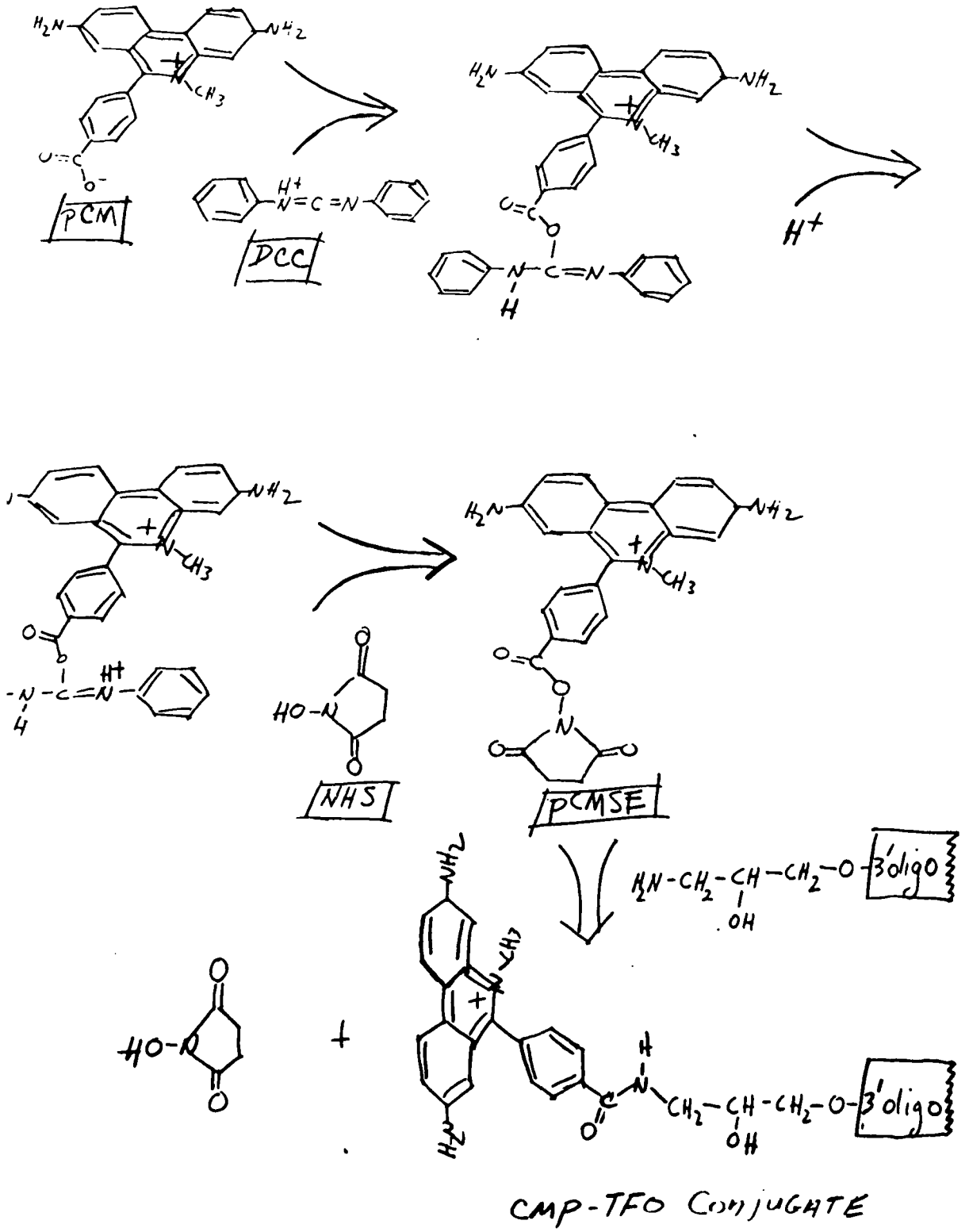
Kdiss = 1x10-7 molar in TFO equivalents.

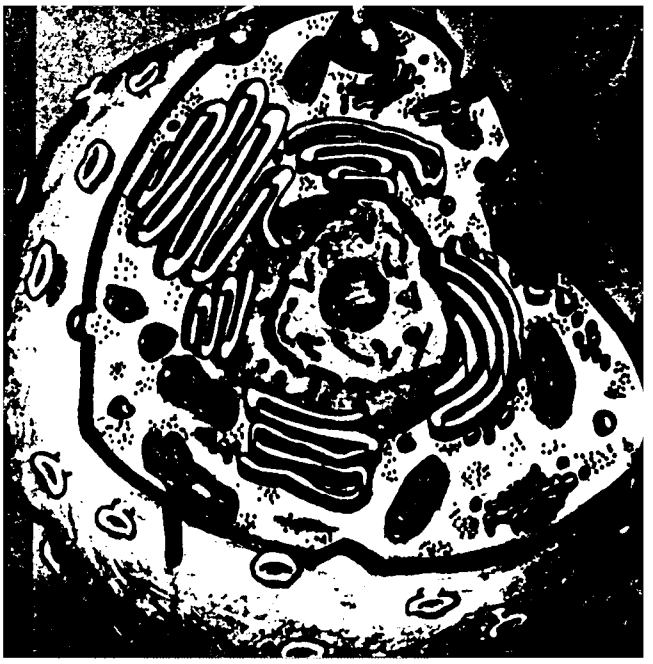
IL28p

Kdiss = 2x10-6 molar.

jure 2. Coupling Chemistry to Form CMP-TFO Conjugates.

We have found that this CMP chemistry can be employed as a » pot" reaction. At reactant concentrations >1mM, the final)ling step appears to be quantitative at a 1:1 dye:TFO ratio, ıssessed by absorbance spectroscopy and by analytical TLC.





35. L.B. Crittenden, G.A. Gulvas, and D.A. Eagen, *Virology* 103, 400 (1980).
36. H.M. Murphy, *Virology* 77, 705 (1977).
37. A.M. Fadly and W. Okazaki, *Poult. Sci.* 61, 1055 (1982).
38. L.B. Crittenden *et al.*, *J. Virology* 61, 722 (1987).

9

The Triplex Technology: A Prospect for Oligonucleotide Therapy at the Gene Level

Correspondence
Michael E. Hogan
Center for Biotechnology
4000Research Forest Drive
The Woodlands, TX 77381

Michael G. Cooney
Department of Biology
Princeton University
Princeton, NJ

The triplex technology employs synthetic oligonucleotide homologues as reagents which bind directly to a duplex DNA target, rather than to single strand mRNA. As such, it can be used as a tool to modulate the transcriptional step of gene expression and provides the potential for targeted DNA chemotherapy. Those distinguishing characteristics emphasize that the triplex and antisense technologies are quite distinct, raising the interesting possibility that the two may be developed into a pair of complementary gene-specific therapies.

Introduction

During the past year, the prospect for oligonucleotide based therapy has become the principle focus of our laboratory at Baylor College of Medicine. However, rather than the design of "antisense" oligonucleotides which bind to mRNA, our program is directed towards the development of modified oligonucleotides which bind directly to specific sites on a DNA helix, thereby forming a three stranded "triplex" which spans a 20-40 base pair (bp) long target site within the control region of a gene.

As for antisense approach, the triplex technology offers the prospect for gene specific therapy. However, because triplex-forming

oligonucleotides are targeted to the gene itself, rather than to the mRNA pool (which comprises a regulated steady state with as many as 10^4 mRNA copies per cell) the two methodologies may prove to have a very different range of application. Specifically, in the (relatively common) instance where the mRNA lifetime is short and the steady state mRNA level is large and well regulated, the antisense technology may be ineffective. In that instance, the triplex technology, with the capacity to repress transcription initiation at the gene itself, may be particularly advantageous.

It is also interesting to consider that, by targeting the helix itself, the triplex technology offers the potential for selective activation of gene expression (by antagonizing a repressor of transcription initiation) and the potential for selective gene destruction (by affixing reactive chemicals to triplex-forming oligonucleotide). An argument can be made that neither of those potentially important applications can be addressed by the antisense technology.

Our interest in site-specific triplex formation is an outgrowth of a preliminary study which we performed in 1986.[1] In that work, we showed that single strand RNA molecules were capable of binding to a discrete site within the control region of the human c-*myc* gene. Although those data were preliminary, it was concluded that the gross secondary structure of the bound complex might be similar to the co-linear triplex formed when polydT binds to the polydA-polydT duplex, a structure that was elucidated 15 years earlier by Arnott and colleagues.[2] In that simple triplex, the oligonucleotide "ligand" (polydT) wraps about the underlying polydA-polydT "target", occupying the major helix groove and engaging in H-bonding with base pairs in the target of the T-A Hoogstein type.[2]

In order to explore the nature of the site-specific binding interaction in c-*myc* we initiated a program in late 1986 of DNA oligonucleotide design: the goal being to determine the "rules" of triplex formation on DNA, using the c-*myc* control region as a model.

Those preliminary observations were published in 1988.[3] In that work, we provided the first evidence that a triplex-forming oligonucleotide could be synthesized with the capacity to bind selectivity to

a target site within a eukaryotic gene promotor under physiological conditions of pH and salt, as assessed by methods which were first developed for the analysis of protein-DNA association: DNA band shift (Figure 1) and DNaseI footprint analysis (Figure 2). Importantly, the c-*myc* target sequence in that instance is not a simple homopolymer (Figure 3).

In order to provide the first experimental assessment of the biological significance of site selective triplex formation, we took advantage of the fact that the chosen target site within the c-*myc* promotor had been shown by Bishop *et al.* to be crucial for the control of c-*myc* transcription initiation.[4] We found that upon administration of an oligonucleotide that forms a site specific triplex with the Bishop trans-activator site, c-*myc* transcription initiation was blocked to completion. The dose response of the effect displayed a midpoint near 5nM of added oligonucleotide (Figure 4), with high apparent specificity, as assessed from a lack of a measurable effect with an oligonucleotide which does not bind to the region.

We consider the c-*myc* effect as the paradigm for a class of synthetic gene repressor, based upon triplex formation at functionally important sites within the control region of a gene.

Our interest in the triplex technology is based upon the hypothesis that triplex-forming oligonucleotides can be developed into a family of therapeutic reagents that exercise their effect by modulating transcription initiation in a gene-specific fashion, *in vivo*, by a mechanism reminiscent of that seen for the c-*myc* gene.

We have focused our attention on gene targets which we consider to be representative of the possible range of application of the technology. In each instance, triplex-forming oligonucleotides have been designed and synthesized against target sites in the transcription control region which was chosen based upon its role as a regulatory site for transcription initiation. For each, the potency of triplex-forming oligonucleotides is being tested with intact cultured cells. For several projects, the program is also being extended to small animal testing. A list of those ongoing projects is presented in Table 1.

In parallel to the Hogan work, both the Dervan and the Helene

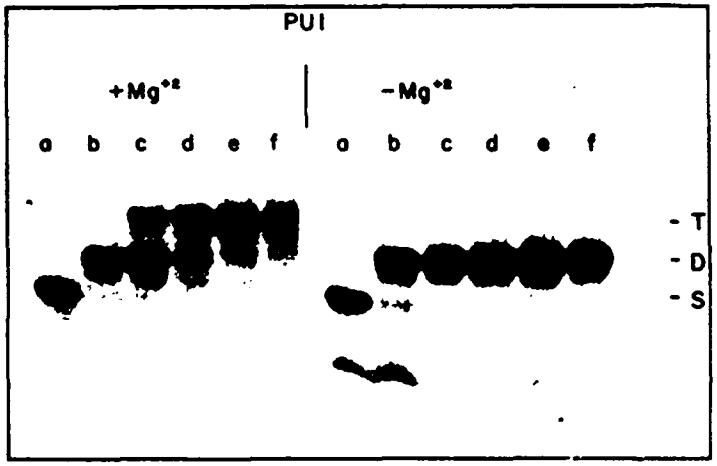
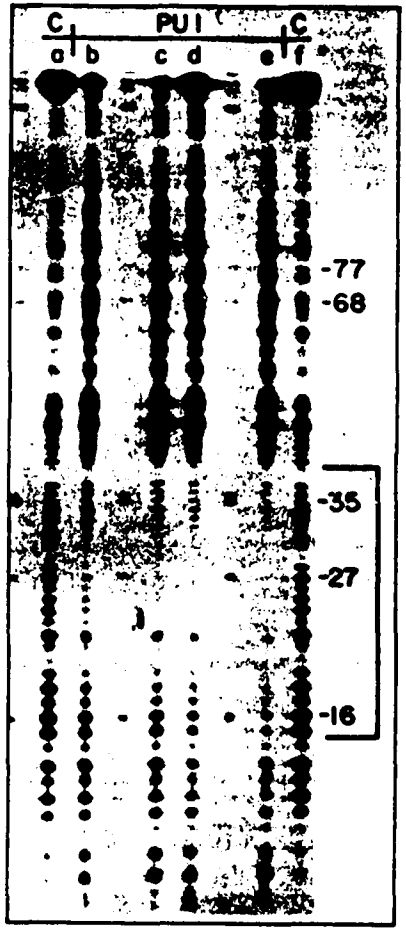


Figure 1. *Detection of the Triplex Formation by the Band Shift Method.* By performing analytical electrophoresis under conditions which stabilize colinear triplex formation (90mM Tris/Borate, 5mM Mg^{2+}, pH 7.8), it is possible to detect single strand oligonucleotides (S), duplex DNA (D) and triplexes (T) as discrete species, based upon characteristic differences in electrophoretic mobility on a 10 percent acrylamide gel. In the data displayed here, the c-*myc* specific oligonucleotide PU1 is shown binding to a 27 bp fragment of the c-*myc gene* (the boxed domain in Figure 1). Lanes b-f correspond to added oligonucleotide concentration of 0, 0.12, 1.2, 12, 24μM. Lane a is a 27-base, single-strand DNA marker. These data also emphasize that in the absence of Mg^{2+} or an equivalent divalent ion (right), triplex formation is too weak to be detected.

laboratories published elegant studies in 1977 and 1988 which have shown that oligonucleotides of the form (CT)n will bind to cloned DNA segments of the form (GA)n-(CT)n at acid pH.[5,6] As was the case for the Hogan studies, those studies showed that triplex formation must be considered as an example of a broader class of site specific ligand-DNA interaction.

It is important to recognize that because those triplexes were based on protonation of C, the type of triplex association described in the Dervan/Helene work are very weak at physiological pH and therefore may be difficult to apply as the basis for a drug therapy. Moreover, the Dervan/Helene studies were performed on simple

Figure 2. *Detection of the Triplex Formation by the DNaseI Footprinting Method.* We have shown that triplex formation prohibits DNaseI cleavage of DNA throughout the duplex DNA target. By analyzing that local cleavage repression, the location of the triplex-forming oligonucleotide binding site can be determined at one base resolution. In this figure, the bracket to the right corresponds to the target site in Figure 1. Lanes a-e correspond to 0, 0.5, 1, 2, 4μM of added PU1.



sequence DNA targets; a finding that has been interpreted by some (including Dervan and Helene) to suggest that potential oligonucleotide binding sites comprise a rather limited subset of the human genome.

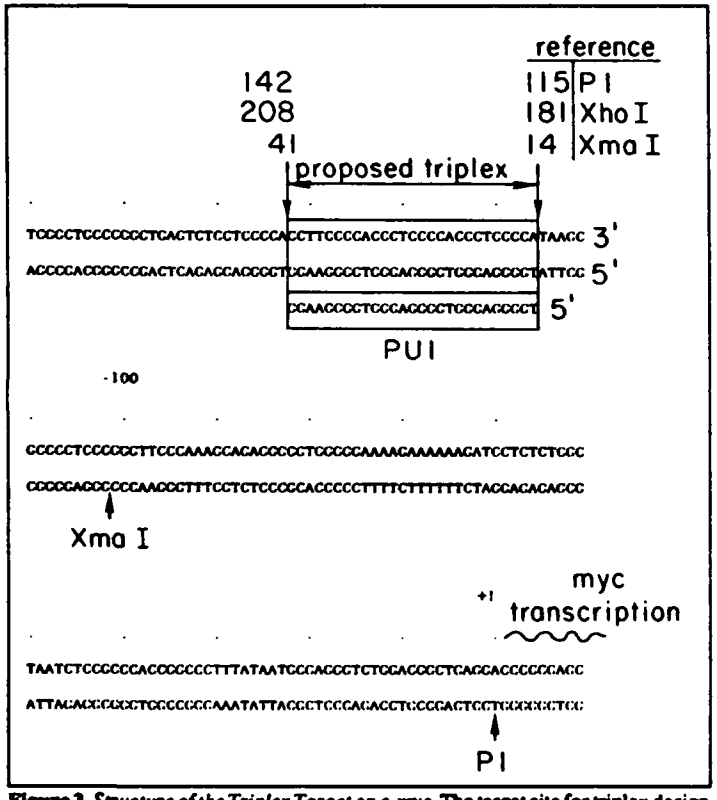


Figure 3. ***Structure of the Triplex Target on c-myc.*** **The target site for triplex design on c-*myc* has been identified, along with one member of the GT27 class of c-*myc* *specific* ligand (PU1). The principle transcription start site (P1) has been identified as +1.**

Working Hypotheses

Two working hypotheses guide our work in this area:

1. By designing triplex forming molecules so that they possess

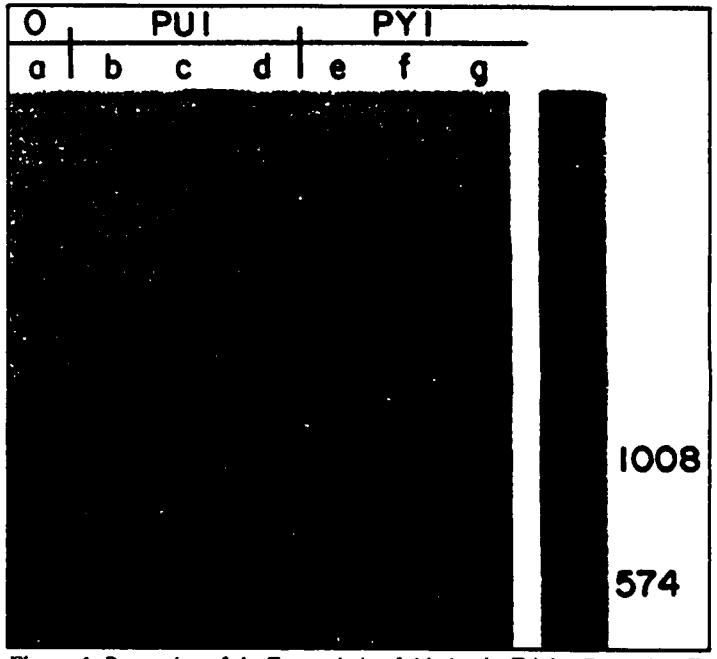


Figure 4. ***Repression of the Transcription Initiation by Triplex Formation.*** **To assess the biological consequences of triplex formation, we have monitored the effect of PU1 binding to the c-*myc* target. Transcription initiation was quantified *in vitro*, employing HeLa cell extracts as the source of RNA polymerase II and other cofactors. The relative transcription rate was assessed by quantifying the amount of radio-labelled RNA resulting from initiation at the P1 start site (resulting in a 1.1kb fragment in this assay). Lane a is a control in the absence of added triplex forming oligonucleotide. Lanes a-d correspond to 0, 100, 66, and 33nM of added PU1. Lanes e-g correspond to 160, 100 and 66nM of added PY1, which does not bind to the target. As seen from the data, PU1 represses transcription initiation with a mid point near to 5nM.**

T-A-T and G-G-C triplets rather than C-G-C triplets, it is possible to achieve stable site-selective triplex formation upon a duplex DNA target at physiological ion concentration, pH, and temperature.

Table 1
Assessment of the Triplex Technology in Cultured Cells

Gene Target Domain	Cell Type in Assay	Assay System	Collaborators (Institution)
1. HIV-1 LTR	monocytes T cells	mRNA, reverse transcriptase	R. Rossen (BCM)
2. CMV immed. early enhancer	hepatocytes	transient assay, lacZ fusion	G. Buffone (BCM)
3. Interleukin 2 receptor	T cells	mRNA, IL2 binding	F. Orson (BCM)
4. c-*myc*	T cells	mRNA, cell growth	F. Orson (BCM) E. Postel (Princeton)
5. mouse insulin	3T3-L1 adipocytes	mRNA, insulin	Triplex Pharmaceutical

2. Triplex formation is not restricted to simple "polypurine" DNA target sites, but is instead a general property of most duplex DNA segments.

Having hypothesized broad applicability of the triplex technology, we believe this class of synthetic gene repressor may have a range of application limited only by a complete knowledge of (i) the rules of triplex formation, (ii) cell uptake properties of oligonucleotides, and (iii) the current understanding of the genetic basis of disease. This class of reagent development has the additional property that because it is based upon the solid phase method of DNA synthesis, a large battery of triplex-forming reagents can be made and tested for site selectivity in the matter of a few days.

In such an experimental environment, we believe that the rate of the triplex design process can approach the limit defined by the designer's capacity to conceive of new structures in response to binding and structural data derived from physical and cellular assay systems.

Investigation of the Rules of Triplex Association

In Figure 5, we have presented binding data derived from the study of triplex-forming oligonucleotide binding to a series of duplex DNA targets. In most instances, data were accumulated by the band-shift method described in an earlier work.[3] Where listed, the data have been confirmed by DNaseI footprinting or other methods. Affinity has been listed as a dissociation constant in units of oligonucleotide concentration.

Our current understanding of the generalizations to be drawn from these studies are listed below as a set of "rules" or operating criteria for triplex design. Apart from the details, it should be noted that the selectivity and overall binding affinity of this class of DNA binding ligand is better, by orders of magnitude, than DNA-specific drugs now in clinical application. Even though there is substantial room for improvement, a good argument can be made that the selectivity of triplex-forming oligonucleotide binding has already begun to approach that displayed by the naturally occurring regulatory proteins.

Binding Specificity

As one estimate of nonspecific binding affinity, we have chosen to measure the binding of triplex-forming oligonucleotides to sheared human DNA. Under standard experimental conditions, we see no evidence for triplex formation for any of the oligonucleotides listed in Figure 5, at the highest accessible duplex binding site concentrations (10^{-2}M in base pair equivalents). In the context of those data, we conclude that triplex-forming oligonucleotides which display specific dissociation constants in the 10^{-7} molar range are binding at least

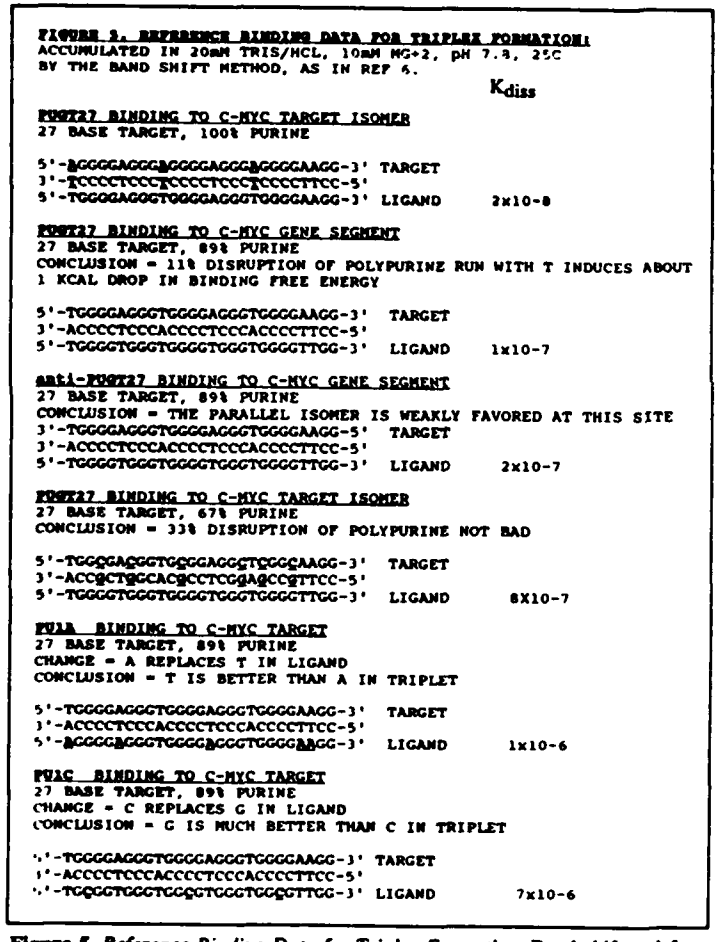


Figure 5. *Reference Binding Data for Triplex Formation.* Band shift and footprinting analysis as in Figures 2 and 3 have been measured as a function of sequence permutation. The resulting data have been fit to the formalism for a simple bimolecular equilibrium. The value for the apparent dissociation constant for triplex formation (K_{diss}) which results from this analysis has been catalogued for each target site complex.

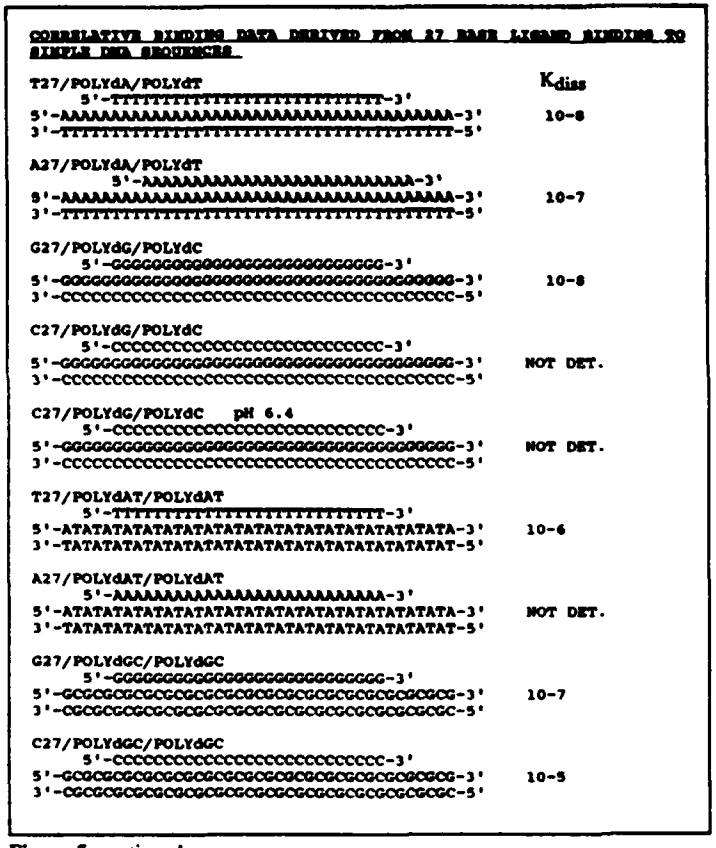


Figure 5, continued.

1000 fold more tightly to their target site than to the integrand of other sequences in the human genome, when assessed by the band shift method.

Rules for the Design of Triplex-forming Oligonucleotides

We have ascertained several trends from these data and from molecular model building (a collaboration with the Institute for Molecular Design, University of Houston). Those trends comprise a set of preliminary design tools. In general, we view such "rules" as primitive, but good enough to design molecules with site selectivity which can be used to manipulate gene expression in cells. We are certain that, as an outgrowth of a broad program of development (including computer assisted molecular modeling), those preliminary rules will be refined in the future, mainly to account for nearest neighbor effects in the duplex binding site, and to allow for components in the triplex-forming reagent other than the standard nucleotide building block. Below, we list our current understanding of the "rules".

Rule 1. Define a binding site domain. In general, the stability of triplex formation increases monotonically with the length of the binding domain. At present, binding sites in the 25-35 base pair range appear to be optimal, based upon consideration of cell uptake, stability, and synthetic considerations.

Rule 2. Define the gross purine/pyrimidine symmetry of the binding site. We define the more purine rich of the two strands of the binding site as the "orienting" strand. As seen from the data above, and from other data obtained in our laboratory (not shown), triplex formation is most stable on targets where the orienting strand is purine rich; however, the rate of stability decrease with respect to decreasing purine content is relatively modest (see above). Several of the targets we have studied in cells fall into a target class which is best described as having an orienting strand with a "purine rich" end abutting a "pyrimidine rich" end. In the context of our current understanding, the triplex which results from oligonucleotide binding to such sites may develop a structural discontinuity at the border between the two subdomains within the target sequence. Consequently, triplex-forming oligonucleotides of that class serve as an interesting test case to be used to understand and optimize binding to such discontinuous targets.

Rule 3. Data at present suggest that a triplex-forming oligonucleotide is best designed so that it can bind throughout the major groove of the target with its backbone parallel to that of the orienting strand. However, the preference may not be large and may itself be dependent upon sequence details of the target. At present, we believe that this aspect of triplex formation is ambiguous enough that we have begun to synthesize both the parallel and antiparallel isomers to any interesting DNA target site.

Rule 4. In either the parallel or antiparallel orientation, place G in the triplex-forming oligonucleotide in opposition to a GC base pair in the target (thereby forming a GGC triplet) and a T in opposition to an AT base pair (to form the TAT) triplet. Such local triplet bonding schemes give rise to H bonding between the triplex-forming oligonucleotide and the purine of the underlying Watson-Crick base pair. Such bonding is likely to be of the Hoogstein or reverse Hoogstein type, depending on sequence context; that is, G=G and T=A bonding respectively. Formation of such "external" bonds occurs through the major groove of the helix, and does not significantly alter the underlying target duplex. Work is in progress in the Hogan laboratory to determine if there are target sequences which produce exceptions to the above simple G=GC and T=AT rule (such as the junctional domain of a bifurcated target site, see above).

Future Prospects

Several laboratories are now actively engaged in the development of triplex-forming oligonucleotides as therapeutic reagents. Based upon that international effort and the similarity between the antisense and triplex technology with respect to cell uptake and stability considerations, it is likely that rapid advancement will be made within the next two years. Below, we list several emerging trends in the area which are likely to mature in the immediate future:

1. The use of nonstandard bases and computer-assisted molecular design. Although Watson-Crick hydrogen bonding between the standard nucleotides is nearly optimal for duplex formation, it is equally likely that nonstandard nucleoside subunits may be better

suited for triplex formation. By pairing computer-assisted molecular design with experimental analysis, it is likely than the triplex technology will make extensive use of substituents of that kind. From that perspective, the triplex technology may prove to be a paradigm for the new era of rational pharmaceutical drug design.

2. Assessment in cell culture. The cellular application of triplex-forming oligonucleotides is now being aggressively investigated in several laboratories. In the coming months, it is likely than gene-specific modulation by these oligonucleotides will be tested with somatic genes of importance and in several virally infected cell lines. It is also likely that data will be available to assess the possibility that the triplex technology can be used to selectively destroy latent viral genomes; for example, to use for the purposes of gene-specific chemotherapy.

3. Assessment in vivo. It is likely that within the next two years, triplex-forming oligonucleotides will be tested for their capacity for selective gene repression *in vivo*. Small animal models are likely candidates, especially transgenics which express human viral antigens. It is also of interest to consider the possibility that triplex-forming oligonucleotides may be tested in a limited clinical context, as topical reagents to manipulate selectively the transcription initiation at gene targets of the skin.

Conclusions

The clinical potential of the triplex technology is difficult to assess at present. However, the promise of the technology is so attractive that research in the area is likely to be vigorous.

References

1. T.C. Joles & M.E. Hogan, *Biochem.* 26, 367 (1987).
2. S. Arnott & E. Selsing, *J. Mol. Biol.* 88, 509 (1974).
3. M. Cooney *et al.*, Science 241, 456 (1988).

4. N. Hay *et al.*, *Genes Dev.* 1,659 (1987).
5. H.E. Moser and P.B. Dervan, *Science* 238, 645 (1987).
6. D. Praseuth *et al.*, *Proc. Natl. Acad. Sci.* 85, 1349 (1987).